



Элементы статистической обработки данных



Статистика — это точная наука, изучающая методы сбора, анализа и обработки данных, которые описывают массовые действия, явления и процессы

Математическая статистика — это раздел математики, изучающий методы сбора, систематизации и обработки результатов наблюдений случайных массовых явлений с целью выявления существующих закономерностей.

Статистика изучает:

- численность отдельных групп населения страны и ее регионов,
- производство и потребление разнообразных видов продукции,
- перевозку грузов и пассажиров различными видами транспорта,
- природные ресурсы и многое другое.

Результаты статистических исследований широко используются для практических и научных выводов.

В настоящее время статистика начинает изучаться уже в средней школе, в ВУЗах это обязательный предмет, потому что связан со многими науками и отраслями.

Чтобы увеличить количество продаж в магазине, чтобы улучшить качество знаний в школе, чтобы двигать страну по экономическому росту, надо проводить статистические исследования и делать соответствующие выводы. И это должен уметь каждый.

Главные цели изучения элементов статистики

- Формирование умений первичной обработки статистических данных;
- изображение и анализ количественной информации, представленной в разных формах (в виде таблиц, диаграмм, графиков реальной зависимостей);
- формирование представлений о важных статистических идеях, а именно: идее оценивания и идее проверки статистических гипотез;
- формирование умений сравнивать вероятности наступления случайных событий с результатами конкретных экспериментов.

Содержание



- Ряд данных
- Объем ряда данных
- Размах ряда данных
- Мода ряда данных
- Медиана ряда
- Среднее арифметическое
- Упорядоченные ряды данных
- Таблица распределения данных
- Подведём итоги
- Номинативный ряд данных
- Частота результата
- Процентная частота
- Группировка данных
- Способы обработки данных
- Подведём итоги

Определение

- **Ряд данных** – это ряд результатов каких-либо измерений.

Например: 1) измерения роста человека
2) Измерения веса человека (животного)
3) Показания счетчика (электроэнергии, воды, тепла...)
4) Результаты в беге на стометровку
И т.д.

Определение

Объемом ряда данных называется количество всех данных.

Например: дан ряд чисел 1; 3; 6; -4; 0
объём его будет равен 5. Почему?

Выполни задание:

В институте сдавали зачет по высшей математике. В группе было 10 человек, и они получили соответствующие оценки: 3, 5, 5, 4, 4, 4, 3, 2, 4, 5.

Определите **объём** данного ряда.

Ответ: 10



Определение

Размах – это разность между наибольшим и наименьшим числами из ряда данных.

Например: если дан ряд чисел 1; 3; 6; -4; 0; 2, то **размах** этого ряда данных будет равен 6 (т.к. $6 - 0 = 6$)

Выполни задание:

В институте сдавали зачет по высшей математике. В группе было 10 человек, и они получили соответствующие оценки: 3, 5, 5, 4, 4, 4, 3, 2, 4, 5.

Определите **размах** данного ряда.

Ответ: 3



Определение

- Модой ряда данных называется число ряда, которое встречается в этом ряду наиболее часто.

Выполни задание:

В институте сдавали зачет по высшей математике. В группе было 10 человек, и они получили соответствующие оценки:

3, 5, 5, 4, 4, 4, 3, 2, 4, 5.

Определите **моду** данного ряда.

Ответ: 4



Определение

- **Медиана с нечётным числом членов** – это число, записанное посередине.
- **Медиана с чётным числом членов** - это среднее арифметическое двух чисел, записанных посередине.

Например: **определить медиану ряда чисел**

1) 6; -4; 5; -2; -3; 3; 3; -2; 3. Ответ: -3

2) -1; 0; 2; 1; -1; 0; 2; -1. Ответ: 0

Выполни задание:

В институте сдавали зачет по высшей математике. В группе было 10 человек, и они получили соответствующие оценки: 3, 5, 5, 4, 4, 4, 3, 2, 4, 5.

Определите **медиану** данного ряда.

Ответ: 4



Определение

- **Среднее арифметическое** - **ЭТО** частное от деления суммы чисел ряда на их количество.

Например: дан ряд чисел -1; 0; 2; 1; -1; 0; 2; -1. Тогда среднее арифметическое будет равно: $(-1+0+2+1+(-1)+0+2+(-1)):8 = 2:8=0,25$

Выполни задание:

В институте сдавали зачет по высшей математике. В группе было 10 человек, и они получили соответствующие оценки: 3, 5, 5, 4, 4, 4, 3, 2, 4, 5.

Определите **среднее арифметическое** данного ряда.

Ответ: 3,9



ПРАКТИЧЕСКАЯ РАБОТА

Задание: охарактеризовать успеваемость ученика Иванова по математике за четвертую четверть.

ВЫПОЛНЕНИЕ РАБОТЫ:

1. Сбор информации:

Выписаны оценки из журнала: 5,4,5,3,3,5,4,4,4.

2. Обработка полученных данных:

объём = 9

размах = $5 - 3 = 2$

мода = 4

медиана = 3

среднее арифметическое = $(5+4+5+3+3+5+4+4+4) : 9 \approx 4$

Характеристика успеваемости: ученик не всегда готов к уроку.

В основном учится на «4». За четверть выходит «4».



Самостоятельно:

Надо найти объём ряда, размах ряда, моду, медиану и среднее арифметическое:

Карточка 1. 22,5; 23; 21,5; 22; 23.

Карточка 2. 6; -4; 5; -2; -3; 3; 3; -2; 3.

Карточка 3. 12,5; 12; 12; 12,5; 13; 12,5; 13.

Карточка 4. -1; 0; 2; 1; -1; 0; 2; -1.

Карточка 5. 125; 130; 124; 131.

Карточка 6. 120; 100; 110.

Определение

- Упорядоченными рядами данных называются ряды, в которых данные расположены по какому то правилу

Как упорядочить ряд чисел? (Записать числа так, чтобы каждое последующее число было не меньше (не больше) предыдущего); или записать некоторые названия «по алфавиту»...

Выполни задание:

Дан ряд чисел:

$-1; -3; -3; -2; 3; 3; 2; 0; 3; 3; -3; -3; 1; 1; -3; -1$

Упорядочить его по возрастанию чисел.

Решение:

$-3; -3; -3; -3; -3; -2; -1; -1; 0; 1; 1; 2; 3; 3; 3; 3$

Получился упорядоченный ряд. Сами данные в нем не изменились, изменился только порядок их следования.



Определение

- **Таблица распределения данных** – это таблица упорядоченного ряда, в котором вместо повторений одного и того же числа записывается количество повторений.

И наоборот, если известна таблица распределения, то можно составить упорядоченный ряд данных.

Например:

Результат измерения	-3	-1	5	7	8
Сколько раз встречается в ряде данных	3	4	2	1	5

Из нее получается такой упорядоченный ряд:

-3;-3;-3;-1;-1;-1;-1;5;5;7;8;8;8;8;8

Выполни задание:

В женском обувном магазине провели статистические исследования и составили соответствующую таблицу по цене обуви и количества продаж:

Цена (руб.):	500	1200	1500	1800	2000	2500
Количество:	8	9	14	15	3	1

Для данных показателей надо найти статистические характеристики:

- составить упорядоченный ряд данных
- объем ряда данных
- размах ряда
- моду ряда
- медиану ряда
- среднее арифметическое ряда данных



И ответить на следующие вопросы:

- Из данных ценовых категорий, обувь за какую цену не следует продавать магазину?
- Обувь, по какой цене следует распространять?
- К какой цене лучше стремиться?



Подведём итоги:

Мы познакомились с **начальными понятиями** того, как происходит **статистическая обработка данных**:

- 1) данные всегда являются результатом какого-либо измерения
- 2) у ряда некоторых данных можно найти:

объём, размах, моду, медиану и

среднее арифметическое

- 3) любой ряд данных можно упорядочить и составить таблицу распределения данных



Определение

- **Номинативный ряд данных** – это НЕ ЧИСЛОВЫЕ ДАННЫЕ, а например, имена; названия; номинации...

Например: список финалистов чемпионатов мира по футболу с 1930 года: Аргентина, Чехословакия, Венгрия, Бразилия, Венгрия, Швеция, Чехословакия, ФРГ, Италия, Нидерланды, Нидерланды, ФРГ, ФРГ, Аргентина, Италия, Бразилия, Германия, Франция

Выполни задание:

Найти из предыдущего примера:

- 1) объём ряда
- 2) моду ряда
- 3) составьте таблицу распределения

Решение: объём = 18; мода – немецкая команда.



Определение

Вероятность случайного события

равна дроби, в знаменателе которой содержится число всех равновероятных возможностей, из которых состоит достоверное событие, а в числителе – число тех возможностей, при которых рассматриваемое событие происходит

Например:



Определение

- **Частота результата** = (сколько раз результат встретился) : (объем данного ряда)

Например: среди 19 данных некоторого измерения один и тот же результат встретился 5 раз. Значит частота данного результата равна 5:19



Определение

Процентная частота = (частота $\cdot$ 100%)

Например: если частота результата равна
 $5:19 = 0,263157\dots$, то процентная частота
будет равна: $0,263 \cdot 100 = 26,3\%$

Часто ответы для процентных частот могут
быть не точными, а приближенными



Определение

- *Группировка данных* – применяется когда различных результатов измерений слишком много. Т.е их объединяют в группы.
- При группировке различных данных информация становится менее точной.

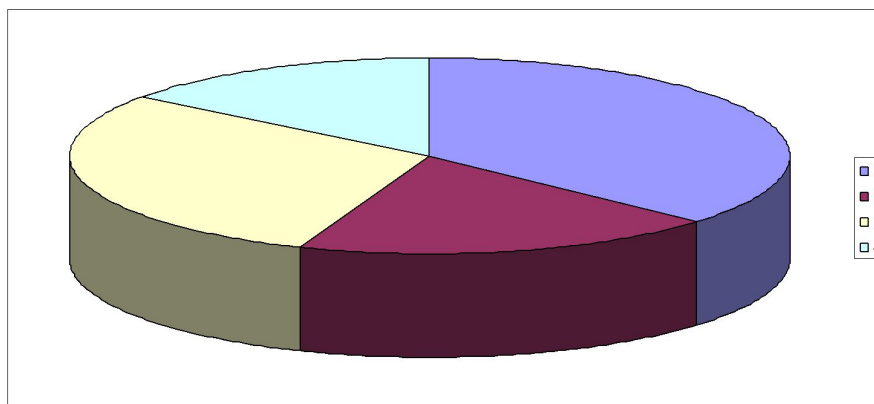


Способы обработки данных:

- Таблица

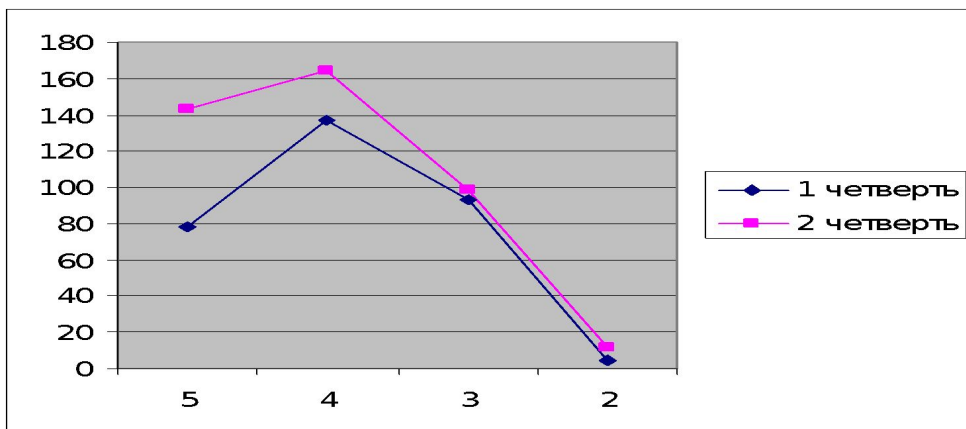
Год обучения	1-4 кл.	5-9 кл	10-11 кл
2007-2008	250	254	80
2008-2009	253	248	78
2009-2010	258	240	73

- Диаграмма круговая (каламбер)

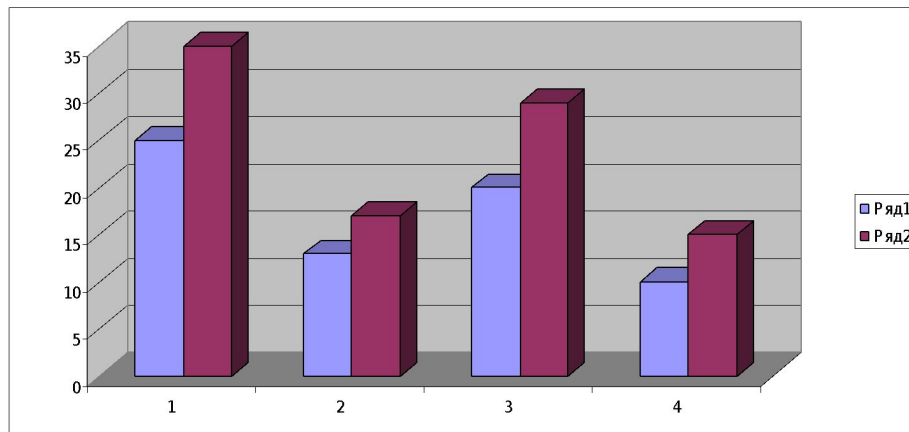


Способы обработки данных

- График

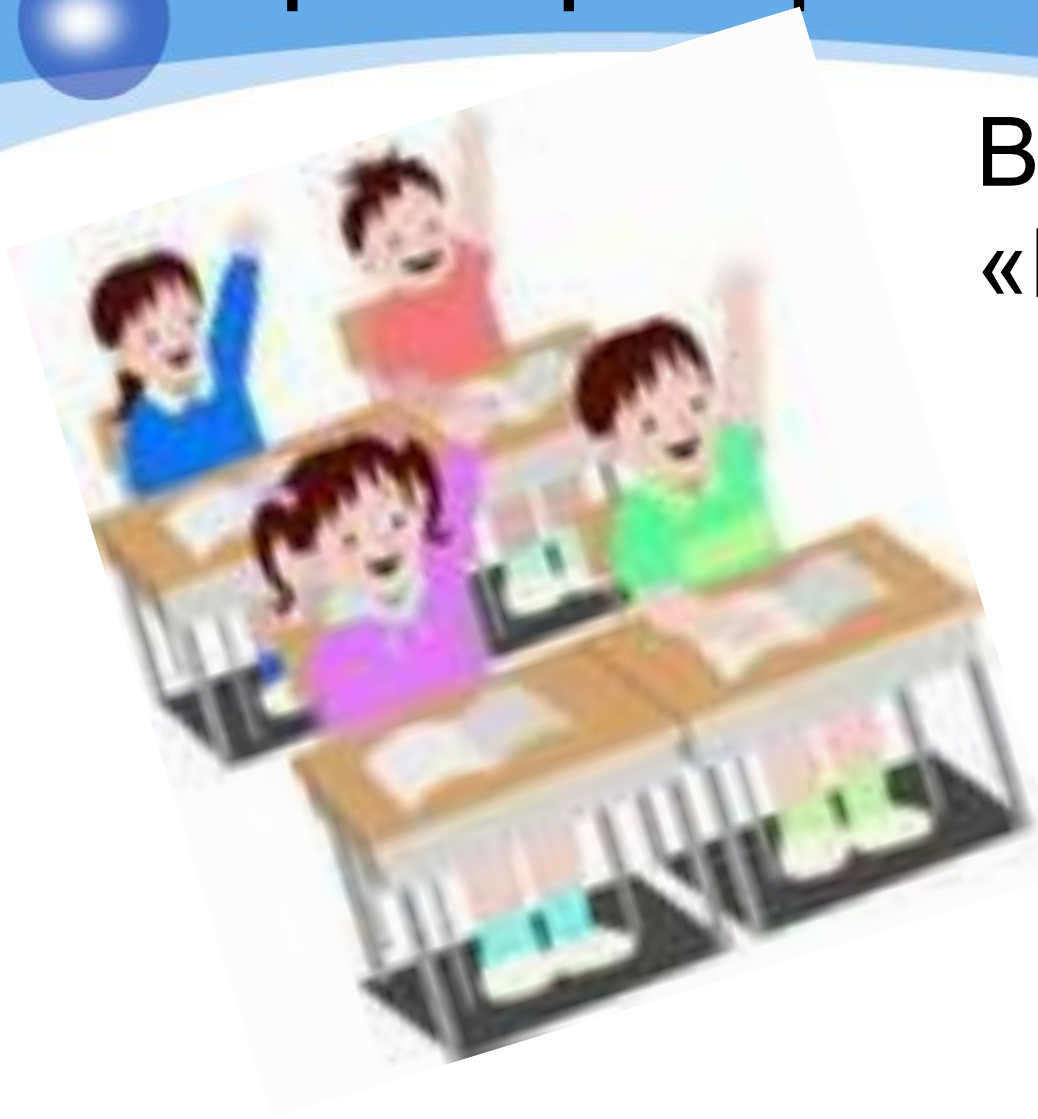


- Гистограмма
(столбчатая диаграмма)



Пример обработки данных:

В финал конкурса
«Мисс
факультета»
вышли 10
студенток,
за которых
голосовали
90 студентов.

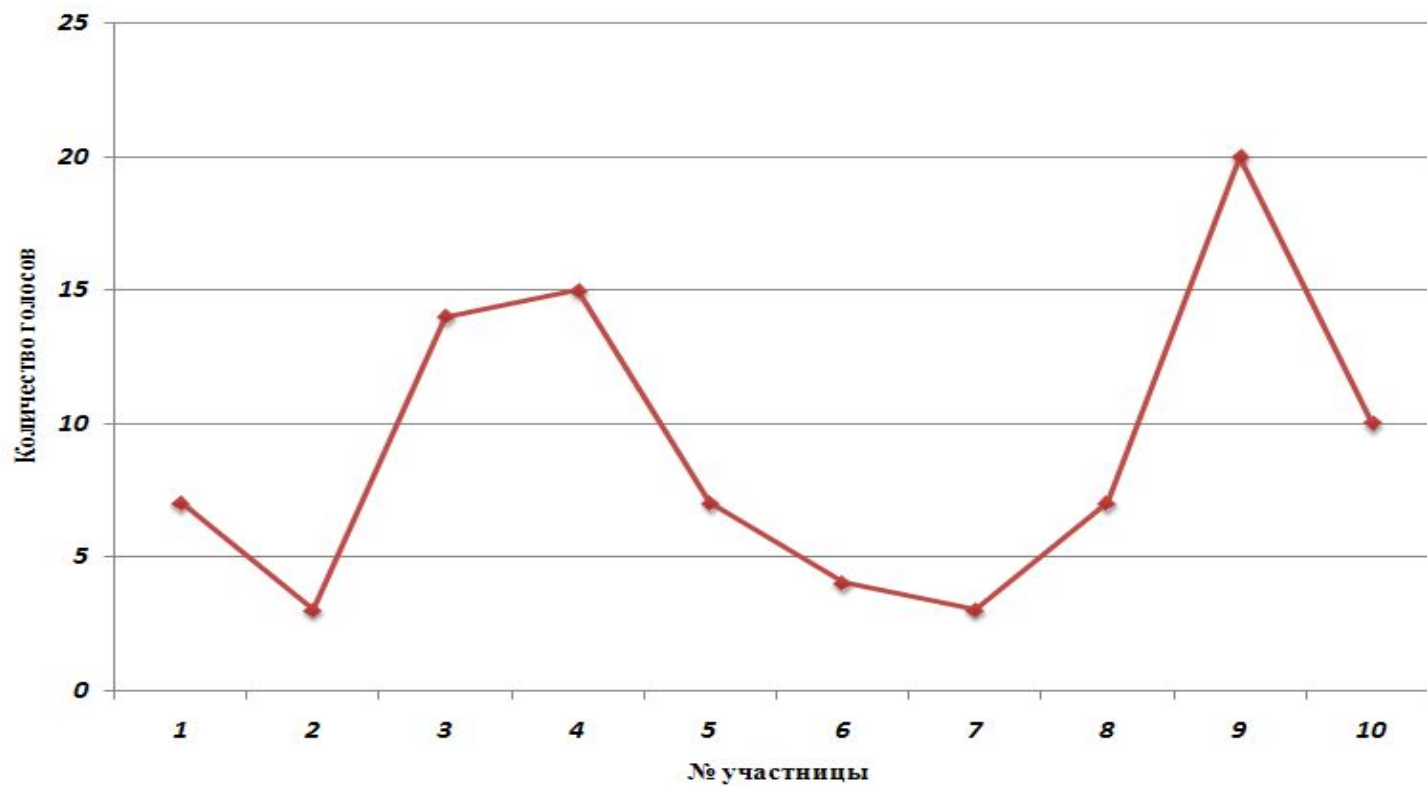


Способы обработки данных

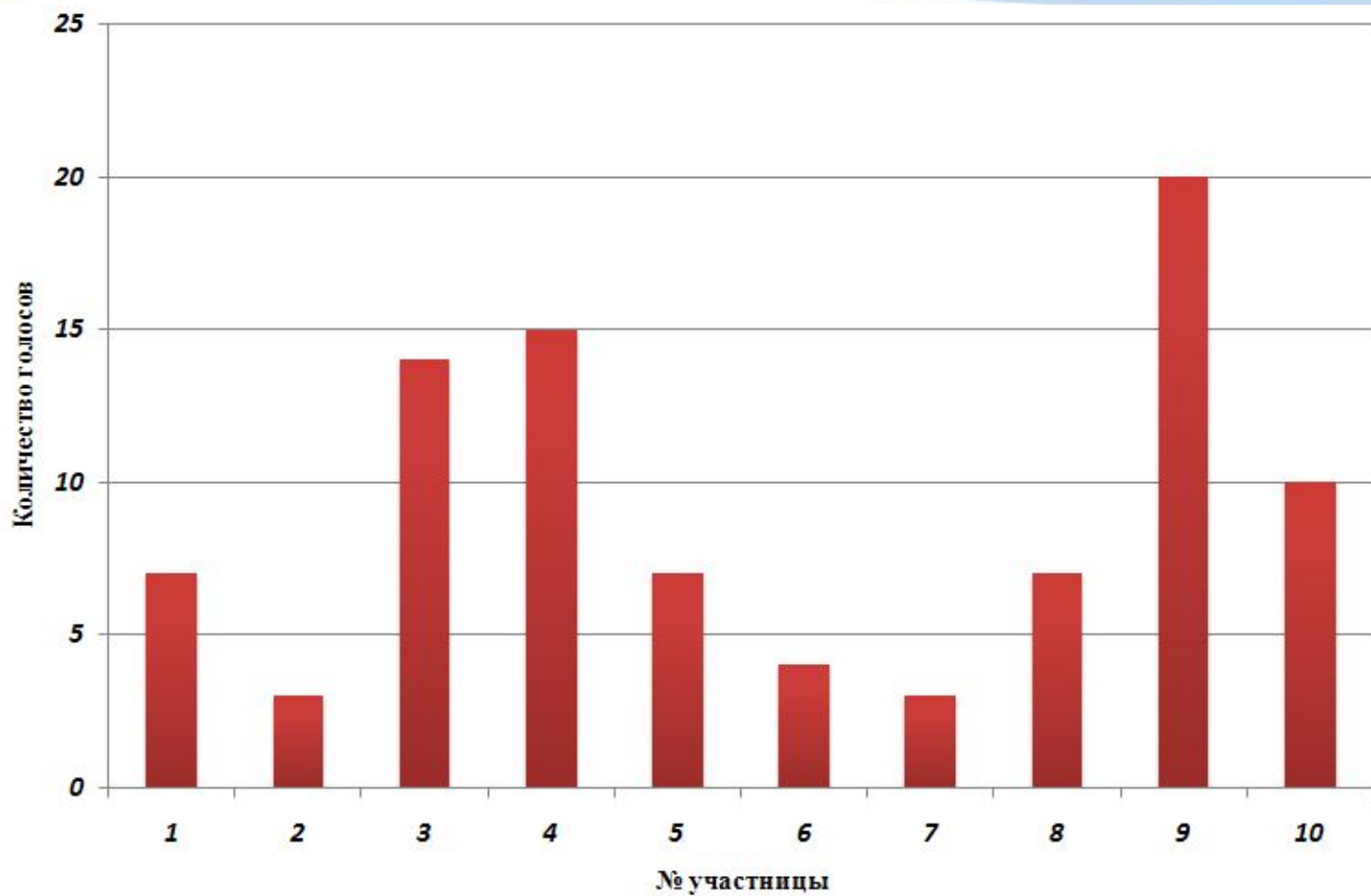
- В таблице приведены результаты голосования за участниц с номерами 1 – 10:

№ участницы	1	2	3	4	5	6	7	8	9	10
Число голосов	7	3	14	15	7	4	3	7	20	10

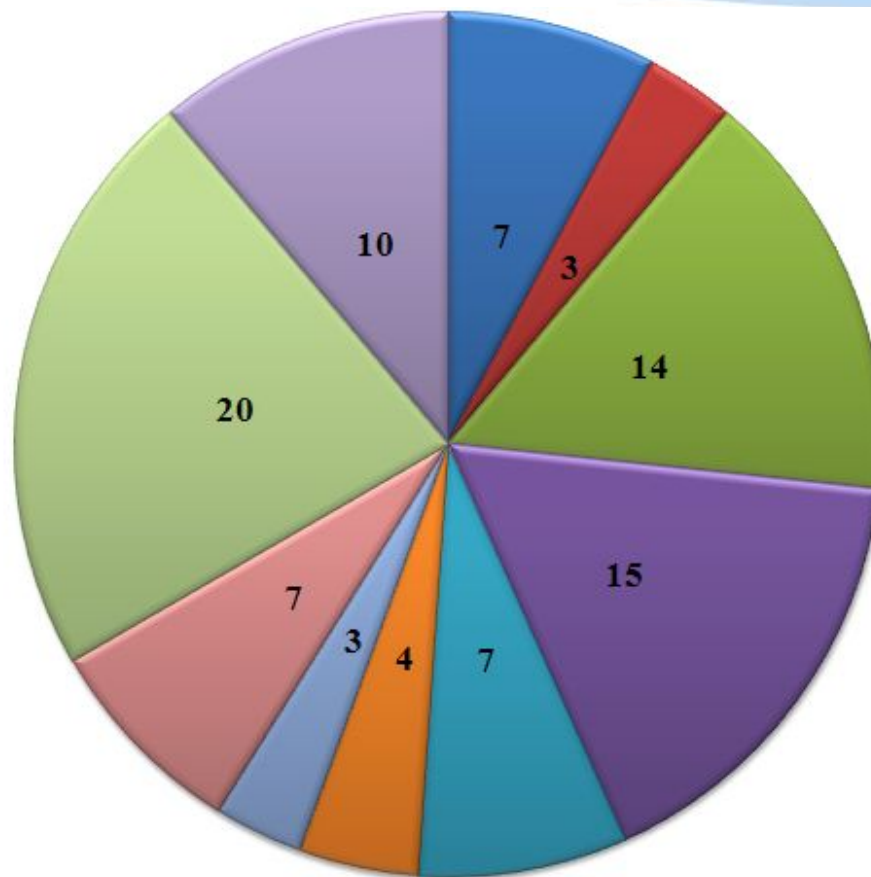
График:



Гистограмма:



Каламбер:



■ 1 ■ 2 ■ 3 ■ 4 ■ 5 ■ 6 ■ 7 ■ 8 ■ 9 ■ 10





Подведём итоги:

Мы узнали, что.....:

- Номинативный ряд – это.....
- Частота результата – это.....
- Группировка данных – это...
- Способы обработки данных:
 - 1) таблица
 - 2) график
 - 3) гистограмма (или столбчатая диаграмма)
 - 4) каламбер (или круговая диаграмма)

Спасибо за внимание!

